May, 1981

LIDS-P-1089

BIFURCATION IN THE PRESENCE OF SMALL NOISE

Shankar Sastry[†]

Omar Hijab[††]

[†] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139. Supported in part by the U.S.Department of Energy under contract DOE-ET-AO1-2295T050.

[††] Division of Applied Sciences, Harvard University, Cambridge, MA 02138. Supported in part by U.S. Office of Naval Research under the Joint Services Electronics Program Contract N00014-75-C-0648.

BIFURCATION IN THE PRESENCE OF SMALL NOISE

The methods of deterministic bifurcation are sensitive to the addition of small amounts of white noise. Thus, for example, in systems whose macroscopic (deterministic) description arises from an aggregation of microscopically fluctuating dynamics the predictions of deterministic bifurcation are incorrect. Here we use Laplace's method of steepest descent to study bifurcation in the presence of small noise. Motivation for this study arises from the Maxwell's equal-area rule for phase transitions in van der Waals gases. The theory is then applied to the study of the dynamics of noisy, constrained or implicitly defined dynamical systems.

Keywords: Bifurcation, Stochastic singular perturbation, jump behavior, Laplace's method of steepest descent, Maxwell's equal area rule.

1. Introduction

Bifurcation is the study of branching in the equilibrium behaviour of a dynamical system in response to small changes in the parameters of the system. Deterministic bifurcation, using as foundations singularity theory, has been fairly successful at explaining a wide variety of such phenomena in fluid mechanics, optics, elastic structures, laser physics and ecology [11]. Nevertheless, the methods of deterministic bifurcation are extremely sensitive to the addition of small amounts of noise. Thus, in systems whose macroscopic description arises from an aggregation of microscopically fluctuating dynamics, thermodynamic systems for example, the predictions of deterministic bifurcation are incorrect. We seek in this paper to set down, to our knowledge for the first time, a theory of stochastic bifurcation - i.e. bifurcation in the presence of small additive noise. We show repeatedly that, in the limit as the intensity of the additive noise tends to zero, the conclusions of this theory are rather different from those of deterministic (or no-noise) bifurcation.

In section 2 we discuss as motivation an example of a thermodynamical phenomenon, isothermic phase transition in Van der Waals gases in which the predictions of deterministic bifurcation theory are incorrect. We then indicate how the addition of small noise predicts the experimentally observed phase transition first studied by Maxwell [9]. In passing we should mention that since the advent of quantum mechanics physicists have been concerned with the derivation of the Van der Waals

equation from first principles (i.e. quantum statistical mechanics). The derivation was first done rigorously in 1963 [7], showing that the Van der Waals equation together with Maxwell's rule are consequences of the quantum theory.

In section 3 we compare deterministic and stochastic bifurcation. We use Laplace's method of steepest descent to compare rigorously the two theories in the limit that the noise intensity goes to zero.

In section 4 we apply the theory of noisy bifurcation to the study of noisy constrained or implicitly defined dynamical systems resulting from the singular perturbation of fast (or 'parasitic') dynamics on some coordinates of the system. The deterministic solution of these systems admit jump discontinuities, including possibly relaxation oscillations, as studied in [12]. The addition of noise however changes the nature of the jump and can in some instances result in the destruction of relaxation oscillations. This is shown explicitly in the case of the degenerate Van der Pol oscillator equation (see for example [15]).

2. Phase Transitions for Van der Waals Gases

One of the models used in the study of phase transition from liquid to gas in thermodynamics is the Van der Waals equation [2] relating the pressure P, the volume V, and the absolute temperature T:

$$rT - (P + \frac{a}{V^2})(V-b) = 0 \qquad (2.1)$$

where a,b,r are positive constants depending on the gas. Loosely speaking, the surface satisfying (2.1) in (P,V,T) space (see figure 1) is a smooth two-dimensional manifold with two fold lines meeting in a cusp at (P_c,V_c,T_c). Isotherms in the (P,V) plane are drawn in figure 2. For temperatures less than T_c the upper left hand corner of figure 2 represents the liquid phase. We study here phase transitions from liquid to gas phase at constant temperature:

2.1 Isothermic Phase Transitions

For temperatures above T_c (supercritical) in figure 2 there is no phase transition (only gas phase). At the critical temperature T_c, the portion of the isotherm to the right (left) of (P_c,V_c) represents the gas (liquid) phase. At subcritical temperatures phase transition is more subtle. If the liquid were allowed to expand 'quasistatically' and isothermally at T_2 by decreasing the pressure, the variation of pressure and volume is described by

$$\dot{P} = f(t) \qquad P(0) = P_o \qquad (2.2)$$

$$g(P,V,T_2) \triangleq rT_2-(P+\frac{a}{V^2})(V-b) = 0 \qquad V(0) = V_o > b \qquad (2.3)$$

where f(t) is a negative function of time. Note that equations (2.2) and (2.3) describe the variation in time of P and V so long as

$$\frac{\partial}{\partial V} g(P,V,T_2) \neq 0 \tag{2.4}$$

since we may then obtain V(t) as a function of P(t) from the implicit function theorem applied to (2.3). At points (P_1,V_1), (P_2,V_2) shown in figure 2, (2.4) is not satisfied and equation (2.3) is singular. A regularization we suggest accounts for the fact that 'quasistatic' expansion of the liquid neglects some 'fast' dynamics (see [12] for details):

$$\dot{P} = f(t) \qquad P(0) = P_o \tag{2.2}$$

$$\varepsilon\dot{V} = g(P,V,T_2) \qquad V(0) = V_o \tag{2.5}$$

The limit of the trajectories of (2.2) (2.5) as $\varepsilon \downarrow 0$ yields a discontinuous change (jump) in volume from (P_1,V_1) to (P_3,V_3) as shown by the dotted line in figure 2. This is the predicted liquid to gas transition. For the converse phase transition choose f(t) to be positive and the jump transition predicted is from (P_2,V_2) to (P_4,V_4).

Thus the predicted deterministic phase transitions are hysteretic (figure 3). Unfortunately, the observed phase transitions are non-hysteretic as shown by the solid line. This line is drawn according to Maxwell's equal area rule [9] - equality of the shaded areas in figure 3.

2.2 Noisy Isothermic Phase Transitions

To explain the observed phase transition we propose to account for noise in equations (2.2), (2.5) stemming from the fact that P, V are aggregations of microscopically stochastic behaviour. In section 4.2 we present rigorous results justifying the manipulations outlined here. We replace (2.2), (2.5) by

$$\dot{P} = f(t) + \sqrt{\mu}\,\xi(t) \qquad P(0) = P_o \tag{2.6}$$

$$\varepsilon\dot{V} = g(P,V,T_2) + \sqrt{\varepsilon\lambda}\,\eta(t) \qquad V(0) = V_o \tag{2.7}$$

where $\xi(\cdot)$ and $\eta(\cdot)$ are standard independent white noises and $\lambda > 0$, $\mu > 0$ scale their variance. For each ε, μ, λ the above equations generate a diffusion $t \to (P(t),V(t))$ in the plane. The evolution of the corresponding probability density $p^{\lambda}_{\mu,\varepsilon}(P,V,t)$ is then given by the Fokker-Planck equation

$$\frac{\partial}{\partial t} p^{\lambda}_{\mu,\varepsilon} = \frac{\mu}{2}\frac{\partial^2}{\partial P^2} p^{\lambda}_{\mu,\varepsilon} - \frac{\partial}{\partial P} p^{\lambda}_{\mu,\varepsilon} f(t) + \frac{\lambda}{2\varepsilon}\frac{\partial^2}{\partial V^2} p^{\lambda}_{\mu,\varepsilon} - \frac{1}{\varepsilon}\frac{\partial}{\partial V} p^{\lambda}_{\mu,\varepsilon} g(P,V,T_2) \tag{2.8}$$

To study (2.8) in the limit $\varepsilon \downarrow 0$, multiply (2.8) by ε and let $\varepsilon \downarrow 0$. Then the limit p^{λ}_{μ} of $p^{\lambda}_{\mu,\varepsilon}$, provided it exists, satisfies

$$\frac{\lambda}{2}\frac{\partial^2}{\partial V^2} p^{\lambda}_{\mu} - \frac{\partial}{\partial V} p^{\lambda}_{\mu} g(P,V,T_2) = 0 \,. \tag{2.9}$$

Solving (2.9) yields

$$p^{\lambda}_{\mu} = k^{\lambda}_{\mu} \exp[-S(P,V,T_2)/\lambda] \tag{2.10}$$

where $k^{\lambda}_{\mu} = k^{\lambda}_{\mu}(P,t)$ and

$$S(P,V,T) = -2\int g(P,V,T)dV$$

$$= PV^2+2a \log V-2PbV+ \frac{2ab}{V} -2rVT_2 .$$

Substituting back in (2.8) and then integrating over V yields

$$\frac{\partial}{\partial t}(k_\mu^\lambda c^\lambda) = \frac{\mu}{2} \frac{\partial^2}{\partial P^2} (k_\mu^\lambda c^\lambda) - \frac{\partial}{\partial P} (k_\mu^\lambda c^\lambda) f(t) \qquad (2.11)$$

where

$$c^\lambda(P) = \int_b^\infty \exp(-S(P,V,T_2)/\lambda)dV < \infty .$$

We see therefore that in the limit $\varepsilon \downarrow 0$ the P process converges to a diffusion whose probability density $k_\mu^\lambda c^\lambda$ satisfies (2.11). Further in the limit $\varepsilon \downarrow 0$, the conditional density $p^\lambda(V|P)$ is given by

$$\frac{1}{c^\lambda(P)} \exp(-S(P,V,T_2)/\lambda) \quad .$$

This is plotted in figure 4 for different values of P. Note that the critical points of $p^\lambda(V|P)$ are exactly the solutions (2.1) with $T = T_2$. For $P > P_2$ there is only one critical point. For $P \leq P_2$ two additional critical points - a local minimum and a local maximum - appear from a <u>fold</u> bifurcation. The new maximum grows in height so that for $P \leq P_4$ it is the global maximum. The old local maximum shrinks and annihilates the minimum again in a <u>fold</u> at P_6. Now Laplace's method of steepest descent (next section) shows that as $\lambda \downarrow 0$ the conditional density $p^\lambda(V|P)$ converges to delta functions supported at the global maxima of $p^\lambda(V|P)$; these 'densities' are plotted in figure 5. The pressure at which the limiting conditional densities jump is the

pressure P_4 at which the two local maxima of $p^{\lambda}(V|P)$ are equal i.e. Maxwell's equal-area rule. Thus the limiting behaviour of (2.6) (2.7) as $\varepsilon \downarrow 0$, $\lambda \downarrow 0$, and $\mu \downarrow 0$, in that order, is the deterministic system

$$\begin{aligned} \dot{P} &= f(t) \\ V &= \psi(P) \end{aligned} \tag{2.12}$$

where ψ is given by figure 5 for $V < V_4$ and $V > \overline{V}_4$ and $\psi(P_4) = V_4$ or $\overline{V}_4$ with probability $\frac{1}{2}$ each.

3.1 Deterministic Bifurcation

Roughly speaking, bifurcation is the study of branching in the equilibrium behaviour of a dynamical system in response to small changes in the parameters of the system. Consider, for example, the class of gradient systems*

$$\dot{x} = -\frac{1}{2} \text{ grad } S(x,u) \tag{3.1}$$

with $S(x,u)$ a smooth (C^∞) proper function growing sufficiently rapidly as $x \to \infty$ in $\mathbb{R}^n$, for each fixed $u_o \in \mathbb{R}^m$. It is well-known [6] that every trajectory of (3.1) converges to an equilibrium point of (3.1) and that every critical point of $S(x,u_o)$ is an equilibrium point of (3.1). Further, if for some u_o, $S(x,u_o)$ is a Morse function [1] then every stable equilibrium of (3.1) is a strict local minimum of $S(x,u_o)$ and conversely. In several practical problems [11] it is of interest to study the variation of the critical points of $S(x,u)$ with the parameter u or, in other words, to study solutions of**

$$D_1 S(x,u) = 0 \tag{3.2}$$

as u varies. If x_o^* is a nondegenerate critical point of $S(x,u_o)$ then there exists a smooth function $x^*(u)$ defined in a neighborhood

---

*Here grad stands for gradient with respect to x.

**$D_1 S(x,u)$ ($D_1^2 S(x,u)$) stands for the first (second) derivative of S with respect to x, while $D_2 g(x,y)$ stands for the first derivative of g with respect to y.

of u_o such that $x^*(u_o) = x_o^*$ and $x^*(u)$ is the only critical point of $S(x,u)$ in some fixed neighborhood of x_o^*. This is the implicit function theorem. Thus smooth continuation of critical points is locally possible from a nondegenerate critical point.

Now suppose x_o^* is a degenerate critical point i.e.

$$\text{rank } D_1^2 S(x_o^*, u_o) = r < n \quad . \tag{3.3}$$

By using the implicit function theorem on the r-dimensional range space of $D_1^2 S(x_o^*, u_o)$ one can show (the method of Lyapunov-Schmidt [12]) that the study of the n equations in n unknowns (3.2) reduces to the study of n-r equations in n-r unknowns

$$N(q,u) = 0 \; ; \tag{3.4}$$

here $q = Px$ where P is the projection onto the kernel of $D_1^2 S(x_o^*, u_o)$ and the first derivative of the bifurcation function N vanishes at (q_o^*, u_o) where $q_o^* = Px_o^*$. The nature of the solution set of (3.2) is thus dependent on the function N. For example suppose r = n-1 (the codimension one case); the function N is then a scalar function of a scalar variable that is at least quadratic near q_o^*. If

$$\frac{\partial^2}{\partial q^2} N(q_o^*, u_o) \neq 0$$

then in a sufficiently small neighborhood of (q_o^*, u_o) there is a unique $q^*(u)$ such that

$$\frac{\partial}{\partial q} N(q^*(u), u) = 0 \quad ,$$

i.e. N has a local maximum or minimum at $(q^*(u),u)$; for definiteness assume that it is a maximum. It can then be shown [4] that if

$$\xi(u) = N(q^*(u),u)$$

then the equations (3.2)

(i) have no solution if $\xi(u) > 0$

(ii) have one solution if $\xi(u) = 0$

(iii) have two solutions if $\xi(u) < 0$,

in a neighborhood of u_o. This is the fold bifurcation and is visualized in figure 6. If

$$\frac{\partial^2}{\partial q^2} N(q_o^*,u_o) = 0 \quad \text{but} \quad \frac{\partial^3}{\partial q^3} N(q_o^*,u_o) \neq 0$$

then $N(q,u_o)$ is at least cubic in $q-q_o^*$ and the bifurcation is a cusp bifurcation (figure 7). Equations (3.2) then have one, two, or three solutions in a neighborhood of u_o.

We do not discuss other bifurcations here. Suffice it to say that the normal form theorems of singularity theory (see Thom [14], Hale [4]) yield universal unfoldings of singularities; the bifurcations of N(q,u) are in general sections of one of these unfoldings.

3.2 Bifurcation in the Presence of Small Noise

Consider, for example, equation (3.1) with added white noise

$$\dot{x} = -\frac{1}{2} \text{ grad } S(x,u) + \sqrt{\lambda}\, \xi(t) \tag{3.5}$$

where $\xi(\cdot)$ is an n-dimensional standard white noise process and λ is a positive constant scaling its variance. For each u in $\mathbb{R}^m$, equation

(3.5) generates a diffusion whose probability density $p^\lambda(t,x,u)$ satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial t} p^\lambda = \frac{1}{2} \sum_{i=1}^{n} [\lambda \frac{\partial^2}{\partial x_i^2} + \frac{\partial}{\partial x_i} (\text{grad } S)_i] p^\lambda \tag{3.6}$$

where $(\text{grad } S)_i$ is the ith component of the vector grad S. We now assume that the derivatives of S grow rapidly enough at ∞ such that as $t \uparrow \infty$, the density $p^\lambda(t,x,u)$ converges exponentially [10] to a unique invariant density $\bar{p}^\lambda(x,u)$; the density $\bar{p}^\lambda$ is then given by

$$\bar{p}^\lambda(x,u) = \exp[-S(x,u)/\lambda]/c^\lambda(u) \tag{3.7}$$

where $c^\lambda(u)$ is chosen such that $\bar{p}^\lambda$ integrates to 1.

Note that for all $\lambda > 0$ and u in $\mathbb{R}^m$ the critical points of $\bar{p}^\lambda(x,u)$ are precisely the equilibrium points of the deterministic system (3.1). Further, if for some u_o, $S(x,u_o)$ is a Morse function then for all $\lambda > 0$ every local maximum of $\bar{p}^\lambda(x,u)$ is a stable equilibrium of (3.1). Thus the study of the bifurcations of the critical points of S(x,u) also yields the bifurcations of the critical points of $\bar{p}^\lambda(x,u)$. Here we study the bifurcations of $\bar{p}^\lambda(x,u)$ in the limit as $\lambda \downarrow 0$ using Laplace's method of steepest descent (for more information see chapter 4 of [5]). We will need the following version of the method.

Theorem. Let, for each $u \in \mathbb{R}^m$, S(x,u) have global minima at $x_1^*(u), \ldots, x_N^*(u)$, where N may depend on u. Let them all be nondegenerate. Let S(x,u) have at least quadratic growth as $x \to \infty$. Then in the limit as $\lambda \downarrow 0$, $\bar{p}^\lambda(x,u)$ converges to

$$\sum_{i=1}^{N} a_i \delta(x-x_i^*) / \sum_{i=1}^{N} a_i$$

where $a_i(u) = \det(D_1^2 S(x_i^*,u))^{-\frac{1}{2}}$. More precisely if $\phi(x,u)$ is a smooth function having polynomial growth as $x \to \infty$, then

$$\lim_{\lambda \downarrow 0} \bar{\phi}_\lambda(u) \equiv \lim_{\lambda \downarrow 0} \int_{\mathbb{R}^n} \phi(x,u) \bar{p}^{\lambda}(x,u)dx = \sum_{i=1}^{N} a_i \phi(x_i^*,u) / \sum_{i=1}^{N} a_i \equiv \bar{\phi}_o(u) .$$

Moreover if the above growth conditions on S and ϕ are uniform in u, for $|u| \leq R$, then $\bar{\phi}_\lambda$ is bounded on $|u| \leq R$ uniformly in $\lambda > 0$ and

$$\int_{|u| \leq R} |\bar{\phi}_\lambda(u) - \bar{\phi}_o(u)|^p du \to 0 \quad ,$$

for all $R > 0$ and $p \geq 1$.

Proof. The proof that $\bar{\phi}_\lambda(u) \to \bar{\phi}_o(u)$ pointwise in u is a slight modification of that appearing in [5]. For the L^p convergence, use the dominated convergence and Egoroff's theorems. □

We therefore see that bifurcations in the presence of small noise are obtained from the study of changes in (non-degenerate) global minima of S(x,u) as a function of u. In contrast to section 3.1, appearance and disappearance of global minima will be points of bifurcation in this context. We close this section with the remark that if the drift in equation (3.5) were not a gradient, then although the invariant density $\bar{p}^{\lambda}(x,u)$ will not necessarily be of the form (3.7), it will be so asymptotically as $\lambda \downarrow 0$, which is enough for the above theorem. This will be developed elsewhere.

4. Noisy Constrained Dynamical Systems

4.1 Deterministic Constrained Dy amical Systems

Consider a constrained or implicitly defined dynamical system of the form:

$$\dot{x} = f(x,y) \tag{4.1}$$

$$0 = g(x,y) \tag{4.2}$$

Here $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, f and g are smooth maps $\mathbb{R}^n \times \mathbb{R}^m$ into $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Assume that 0 is a regular value of g. We interpret (4.1), (4.2) as describing implicitly a dynamical system on the n-dimensional manifold

$$M = \{(x,y) \mid g(x,y) = 0\} \subset \mathbb{R}^{n+m} \tag{4.3}$$

Alternatively, one can think of (4.1) as describing a control system with control variable y and (4.2) describing implicitly a feedback control law, a situation that arises frequently in optimal control for example; (4.2) would then correspond to the Euler-L agrange equations. The vector field X(x,y) on M is defined by specifying its projection along the x-axis, namely

$$\pi X(x,y) = f(x,y) \ , \tag{4.4}$$

where π is the projection map $(x,y) \to x$. At points where the mxm matrix $D_2 g(x,y)$ is full rank, it is clear that (4.4) uniquely specifies X(x,y). Difficulties occur at points (x,y) where $D_2 g(x,y)$ is not of full rank and f(x,y) is transverse to $\pi TM(x,y)$.

It would seem then that the trajectory should instantaneously jump off the manifold from (x,y) to some other point (x,y') on M (x is constrained to vary continuously by (4.1)). (see figure 8). This intuition is made precise in [12] by interpreting solution trajectories of (4.1), (4.2) to be the 'degenerate' limit as $\varepsilon \downarrow 0$ of the trajectories of

$$\dot{x} = f(x,y) \tag{4.1}$$

$$\varepsilon\dot{y} = g(x,y) \ , \tag{4.5}$$

provided the limits exist. The resulting solution concept including jump discontinuities and applications are developed in [12]. We illustrate the theory with an example.

Consider the system of equations in the plane given by

$$\dot{x} = y \tag{4.6}$$

$$0 = -x-y^3+y \ , \tag{4.7}$$

the degenerate Van der Pol oscillator [15]. The phase portrait for the degenerate system including jumps from two fold singularities of the projection map $\pi : (x,y) \to x$ is shown in figure 9 . Note the relaxation oscillation formed by including the two jumps.

4.2 Noisy Constrained Systems

We add noise to the system (4.1), (4.5) to obtain

$$\dot{x} = f(x,y) + \sqrt{\mu}\ \xi(t) \qquad x(0) = x_o \tag{4.8}$$

$$\varepsilon\dot{y} = g(x,y) + \sqrt{\lambda\varepsilon}\ \eta(t) \tag{4.9}$$

Here $\xi(\cdot)$ and $\eta(\cdot)$ are, as before, independent $\mathbb{R}^n$-valued and $\mathbb{R}^m$-valued white noise processes and λ, μ scale their variance. For each $\mu, \lambda > 0$ systems of this kind have been studied in the limit $\varepsilon \downarrow 0$ extensively by Papanicolaou, Strook and Varadhan [10]. Our contribution here is to study the behaviour of (4.8) (4.9) in the further limit $\lambda \downarrow 0$ followed by $\mu \downarrow 0$, and to present it in contrast to the results of section 4.1.

In order to apply Laplace's method, restrict attention to the case

$$g(x,y) = -\frac{1}{2} \text{ grad } S(x,y) ,$$

the gradient taken with respect to y, for some S smooth proper and growing sufficiently rapidly as $y \to \infty$, uniformly for $|x| \leq R$, for all $R > 0$. Note that here x and y play the roles of u and x in section 3.2. For each $\varepsilon, \lambda, \mu > 0$, the evolution of the corresponding probability density $p^{\lambda}_{\mu,\varepsilon}$ is governed by

$$\frac{\partial}{\partial t} p^{\lambda}_{\mu,\varepsilon} = (L_o^* + \frac{1}{\varepsilon} L_1^*) p^{\lambda}_{\mu,\varepsilon}$$

where L_o^*, L_1^* are the formal adjoints of the operators L_o, L_1 given by

$$L_o p = \frac{1}{2} \sum_{i=1}^{n} [\mu \frac{\partial^2}{\partial x_i^2} + f_i \cdot \frac{\partial}{\partial x_i}] p$$

and

$$L_1 p = \frac{1}{2} \sum_{i=1}^{m} [\lambda \frac{\partial^2}{\partial y_i^2} + g_i \frac{\partial}{\partial y_i}] p \quad .$$

The conditional density of y given x is, in the limit $\varepsilon \downarrow 0$, given by

$$\bar{p}^{\lambda} = \exp[-S(x,y)/\lambda]/c^{\lambda}(x)$$

where $c^{\lambda}(x)$ is chosen so that $\bar{p}^{\lambda}$ integrates to one. Set

$$\bar{f}_{\lambda}(x) = \int_{\mathbb{R}^m} f(x,y)\bar{p}^{\lambda}(x,y)dy \quad ,$$

for $\lambda > 0$. Assume that $f(x,y)$ has at most polynomial growth as $y \to \infty$, uniformly in $|x| \leq R$, for all $R > 0$. It then follows that $\bar{f}_{\lambda}(x)$ is bounded on $|x| \leq R$ uniformly in $0 < \lambda < \lambda_o$, for some λ_o sufficiently small. We assume in what follows that $\bar{f}_{\lambda}(x)$ has linear growth uniformly in $0 < \lambda \leq \lambda_o$. Set

$$\bar{L}_o p = \frac{1}{2} \sum_{i=1}^{n} [\mu \frac{\partial^2}{\partial x_i^2} + (\bar{f}_{\lambda})_i \frac{\partial}{\partial x_i}]p \quad .$$

The operator $\bar{L}_o$ is the operator L_o averaged over the invariant density $\bar{p}^{\lambda}$ of y given x.

Theorem [10] As $\varepsilon \downarrow 0$ the first component $t \to x(t)$ of the solutions of equations (4.8), (4.9) converges weakly in $C([0,T]; \mathbb{R}^n)$ to the unique diffusion, denoted by $t \to x_{\lambda,\mu}(t)$, governed by $\bar{L}_o$.

We now have the following

Theorem. As $\lambda \downarrow 0$ the diffusion $t \to x_{\lambda,\mu}(t)$ converges weakly in $C([0,T]; \mathbb{R}^n)$ to the unique diffusion $t \to x_{\mu}(t)$ satisfying in law

$$\dot{x} = \bar{f}_o(x) + \sqrt{\mu}\ \xi(t) \quad x(0) = x_o \tag{4.10}$$

where

$$\bar{f}_o(x) = \sum_{i=1}^{N} a_i(x) f(x, y_i^*(x)) / \sum_{i=1}^{N} a_i(x)$$

as in section 3.2, and $y_1^*(x), \ldots, y_N^*(x)$ are the nondegenerate global minima of $S(x,\cdot)$.

Proof. The proof of this and the next theorem mimic that of the previous theorem and so we only outline the proof. The reader may prefer to master the proof of the preceding theorem appearing in [10] first. We first show that the measures on $C([0,T];\mathbb{R}^n)$ corresponding to $t \to x_{\lambda,\mu}(t)$ are relatively weakly compact in $C([0,T];\mathbb{R}^n)$. The second step is to identify any limiting measure of the family $x_{\lambda,\mu}$, $\lambda > 0$, as the unique solution of the martingale problem associated to (4.10). It is essential in both steps 1 and 2 that μ remain fixed and positive.

To prove compactness in $C([0,T];\mathbb{R}^n)$ it suffices to show that

$$\lim_{\delta \downarrow 0} \sup_{\lambda > 0} P\Big(\sup_{\substack{|t-s| \leq \delta \\ 0 \leq t,s \leq T}} |x_{\lambda,\mu}(t) - x_{\lambda,\mu}(s)| \geq \varepsilon \Big) = 0$$

for all $\varepsilon > 0$ ([13] page 351). This follows from the fact that

$$\dot{x}_{\lambda,\mu} = \overline{f}_\lambda(x_{\lambda,\mu}) + \sqrt{\mu}\ \xi(t) \qquad x(0) = x_o$$

in law (which is all we need here), and the assumptions on $\overline{f}_\lambda(x)$ ([3] page 120).

For step 2, take a C^∞ function ϕ whose support is in $|x| < R$ and consider

$$\phi(x_{\lambda,\mu}(t)) - \phi(x_o) - \int_0^t \overline{f}_\lambda(\phi)(x_{\lambda,\mu}(s))ds - \frac{\mu}{2}\int_0^t \Delta(\phi)(x_{\lambda,\mu}(s))ds\ ,$$

where $f(\phi)$ is the directional derivative of ϕ in the direction of the vector field f and Δ is the Laplacian. For each $\lambda > 0$ this expression is a martingale and the idea is that the limit is also a martingale.

Note that to get the correct expression in the limit, one has to show that

$$E\left(\int_0^T \left|\bar{f}_\lambda(\phi)(x_{\lambda,\mu}(s))-\bar{f}_o(\phi)(x_{\lambda,\mu}(s))\right|ds\right) \tag{4.11}$$

goes to zero as $\lambda \downarrow 0$. This follows from the fact that for all p sufficiently large there is a $K > 0$ depending only on p, a uniform bound for $\bar{f}_\lambda(x)$, $|x| \leq R$, $0 < \lambda \leq \lambda_o$, and T, such that (4.11) is bounded by

$$K\left|\left|\bar{f}_\lambda(\phi)-\bar{f}_o(\phi)\right|\right|_p \tag{4.12}$$

where the L^p norm is over $|x| \leq R$ ([8] page 52). But from section 3.2, we know that (4.12) goes to zero as $\lambda \downarrow 0$. This completes the proof.

Finally we can let $\mu \downarrow 0$ to obtain

Theorem. The family $t \to x_\mu(t)$, $\mu > 0$ is relatively weakly compact in $C([0,T]; \mathbb{R}^n)$. Any limiting diffusion of the diffusions $t \to x_\mu(t)$, $\mu > 0$, as $\mu \downarrow 0$ then satisfies the ordinary differential equation

$$\dot{x} = \bar{f}_o(x) \quad , \quad x(0) = x_o \quad . \tag{4.13}$$

Proof. As before one checks that $t \to x_\mu(t)$, $\mu > 0$, is relatively weakly compact, as before any limiting diffusion is then governed by $\bar{f}_o$, i.e.

$$\phi(x(t)) = \phi(x) + \int_0^t \bar{f}_o(\phi)(x(s))ds + \text{martingale} \quad .$$

Since the variance of the martingale is

$$E\int_0^t [\bar{f}_o(\phi^2)-2\phi\bar{f}_o(\phi)](x(s))ds$$

and $f(\phi^2)-2\phi f(\phi)$ is zero for any vector field f, it follows that the martingale part is identically zero and (4.13) holds. □

Of course there may be many solutions to (4.13).

Let us consider, as an example, the noisy version of the degenerate Van der Pol oscillator. Consider

$$\dot{x} = y + \sqrt{\mu}\ \xi(t)$$

$$\varepsilon\dot{y} = -x-y^3+y + \sqrt{\lambda\varepsilon}\ \eta(t)$$

For $\lambda,\mu > 0$ as $\varepsilon \downarrow 0$ the x-process converges to one satisfying

$$\dot{x} = \bar{y}^{\lambda}(x) + \sqrt{\mu}\ \xi(t)$$

where

$$\bar{y}^{\lambda}(x) = \int_{-\infty}^{+\infty} y \exp\frac{2}{\lambda}\left(-xy - \frac{y^4}{4} + \frac{y^2}{2}\right)dy \Big/ \int_{-\infty}^{+\infty} \exp\frac{2}{\lambda}\left(-xy - \frac{y^4}{4} + \frac{y^2}{2}\right)dy \ .$$

In figure 10, $\bar{y}^{\lambda}(x)$ is plotted for $\lambda_1 > \lambda_2 > 0$.

In the further limit that $\lambda \downarrow 0$ followed by $\mu \downarrow 0$, x satisfies

$$\dot{x} = \psi(x) \qquad x \neq 0$$

$$= 0 \qquad x = 0$$

where $\psi(x)$ is shown heavy in figure 10. Note that ψ is discontinuous at $x = 0$, since the support of the conditional density jumps from one leg of the curve $x = y^3-y$ to the other leg as shown in figure 11. Consequently the relaxation oscillation is broken up by the presence of small noise.

Acknowledgements:

The authors would like to thank D. Castanon, A. Willsky, B. Levy, and M. Coderch for many helpful discussions.

References

[1] R. Abraham and J. Marsden, Foundation of Mechanics, 2nd ed., Addison-Wesley, Reading, MA 1978.

[2] H.B. Callen, Thermodynamics, Wiley, New York, 1960.

[3] W. Fleming and R. Rishel, Deterministic and Stochastic Optimal Control, Springer-Verlag, N.Y. 1975.

[4] J.K. Hale, "Generic Bifurcation with Applications," in Nonlinear Analysis and Mechanics, Heriot-Watt Symposium, v. 1, Pitman 1977.

[5] O. Hijab, "Minimum Energy Estimation," Ph.D. Dissertation, University of California, Berkeley, December 1980.

[6] M.W. Hirsch and S. Smale, Differential Equations, Dynamical Systems, and Linear Algebra, Academic Press, 1975.

[7] M. Kac, G.E. Uhlenbeck, P.C. Hemmer, "On the Van der Waals Theory of the Vapor-Liquid Equilibrium," Journal of Mathematical Physics 4, #2, February 1963, pp. 216-247.

[8] N.V. Krylov, Controlled Diffusion Processes, Springer-Verlag, N.Y., 1980.

[9] J.C. Maxwell, "On the Dynamical Evidence of the Molecular Constitution of Bodies," Nature 11, (1875), pp. 357-374.

[10] G.C. Papanicolaou, D. Strook, S.R.S. Varadhan, "Martingale Approach to Some Limit Theorems," in Statistical Mechanics, Dynamical Systems, and the Duke Turbulence Conference, ed. D. Ruelle, Duke University Math. Series, 3, Durham, N.C. 1977.

[11] T. Poston and I.W. Stewart, Catastrophe Theory and Its Applications, Pitman, 1978.

[12] S.S. Sastry, C.A. Desoer, P.P. Varaiya, "Jump Behaviour of Circuits and Systems," ERL Memo No. M80/74, University of California Berkeley (To appear IEEE Trans. Circuits and Systems).

[13] D. Strook and S.R.S. Varadhan, "Diffusion Processes with Continuous Coefficients, I", Comm. Pure and Applied Math. XXII, pp. 345-400 (1969).

[14] R. Thom, Structural Stability and Morphogenesis, Benjamin 1975.

[15] E.C. Zeeman, "Differential Equations for the Heart Beat and Nerve Impulse" in Towards a Theoretical Biology 4, ed. C.H. Waddington, Edinburgh University Press 1972.

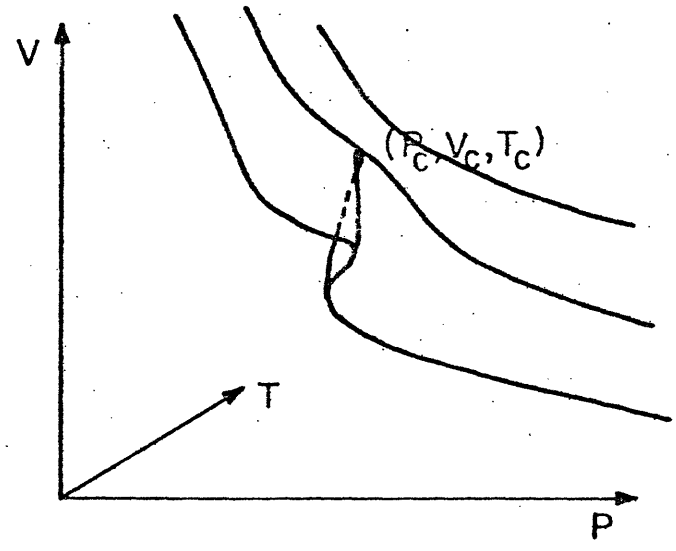

Figure 1. The Surface Satisfying (2.1) Showing a Cusp point at (P_c, V_c, T_c)

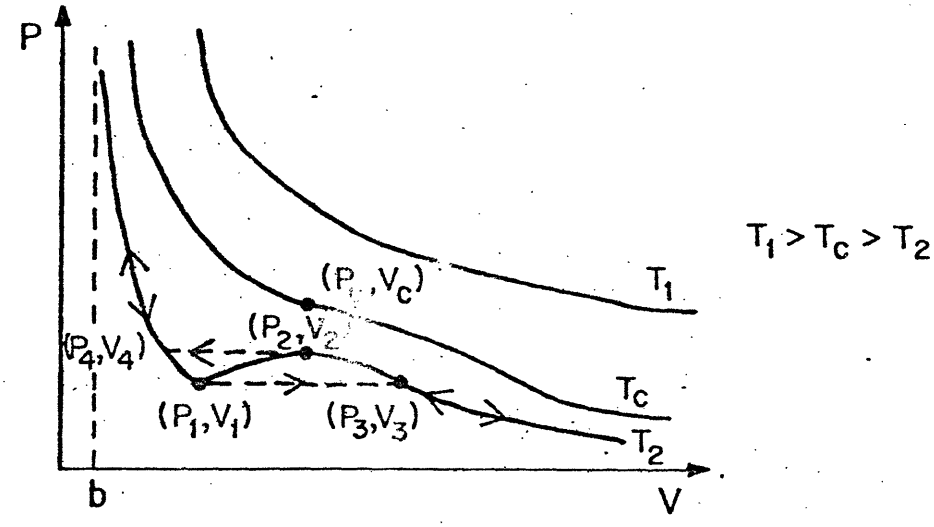

Figure 2. Isotherms of the Van der Waals Gas Equation

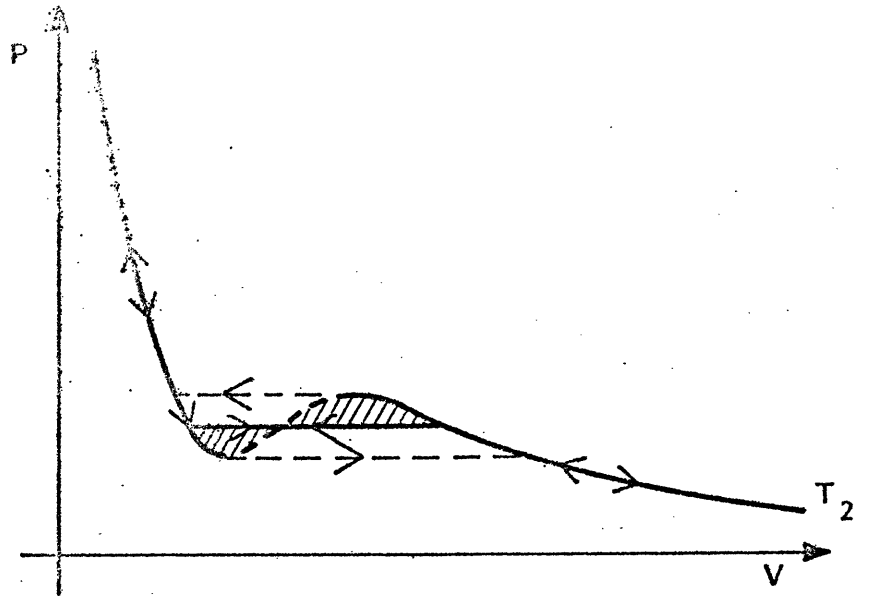

Figure 3. Showing the Discrepancy Between the Deterministic Prediction (Dotted) and Experimentally Observed (Solid) Phase Transition

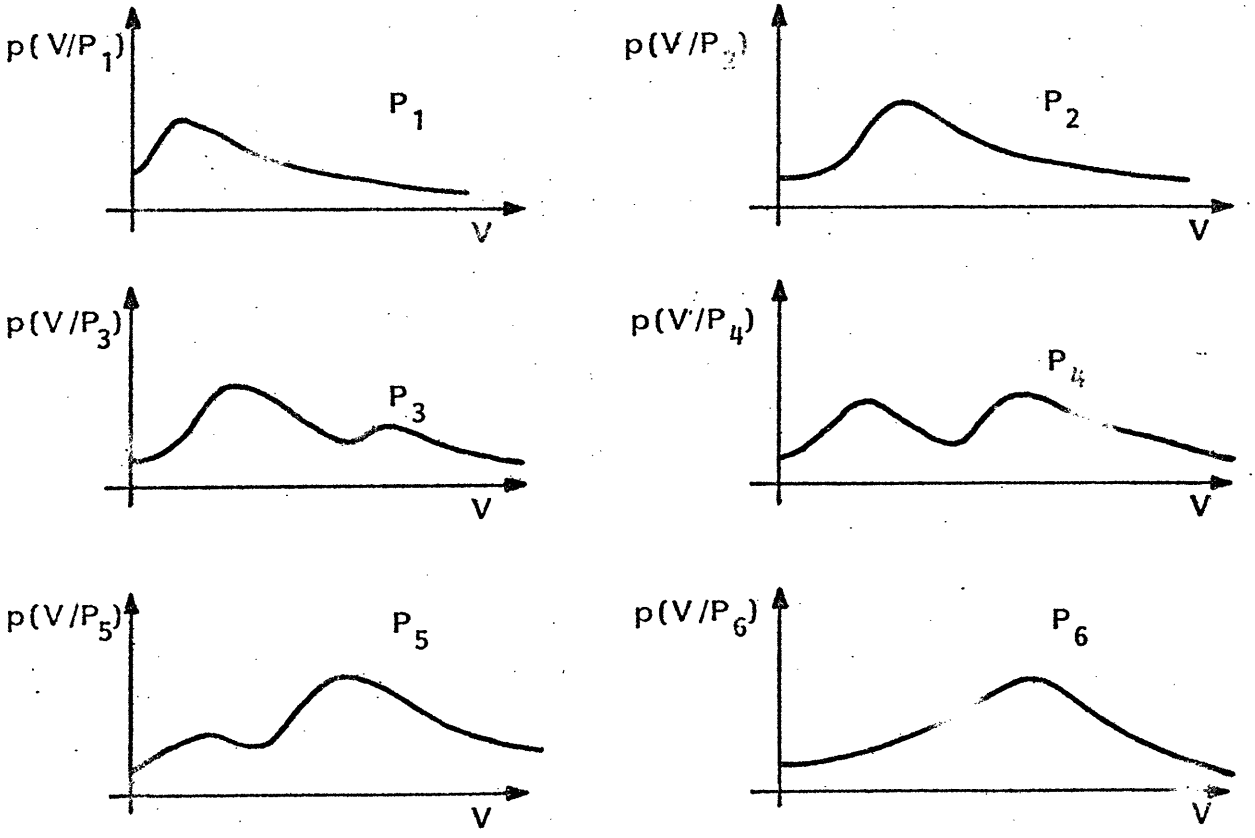

Figure 4. The Conditional Density p(V/P) Plotted for Decreasing Values of Pressure $P_1 > P_2 > P_3 > P_4 > P_5 > P_6$.

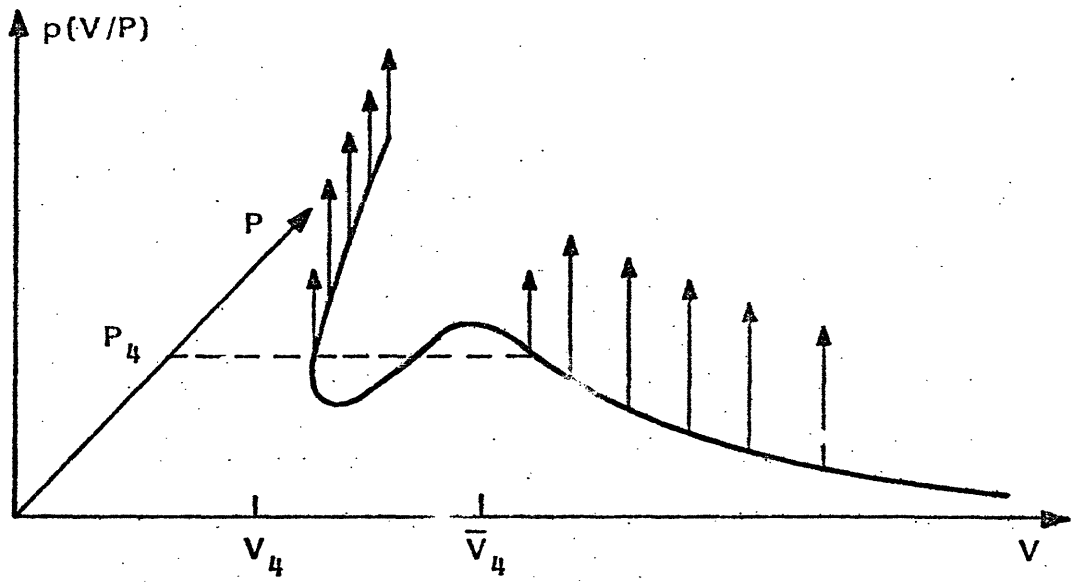

Figure 5. The Limit of the Conditional Density p(V/P) when $\lambda \downarrow 0$

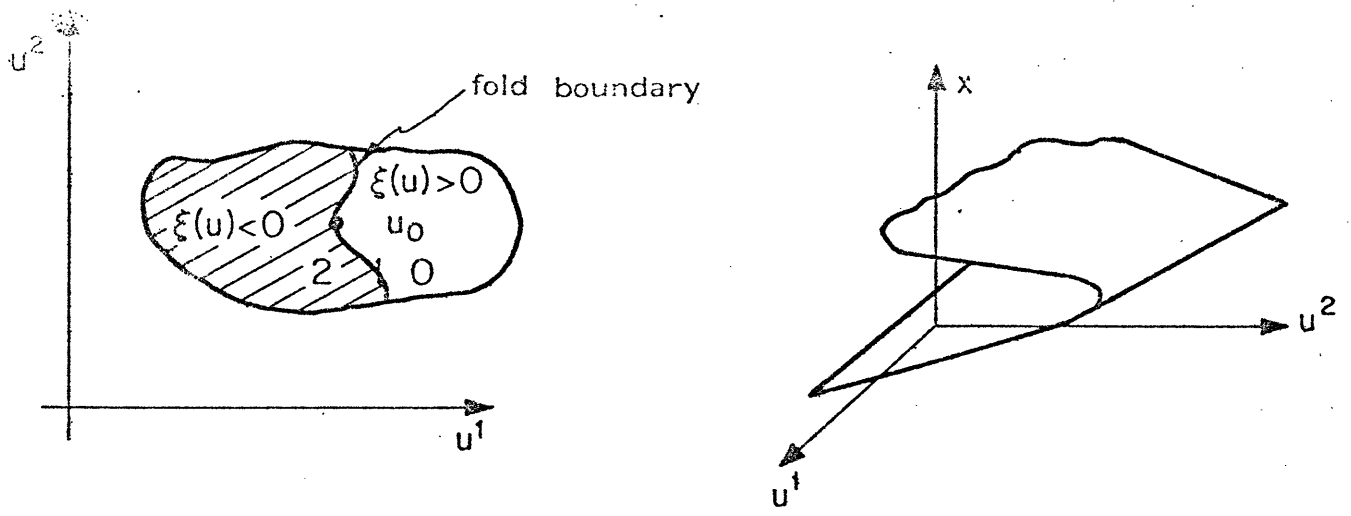

Figure 6. Visualisation of the Fold Bifurcation

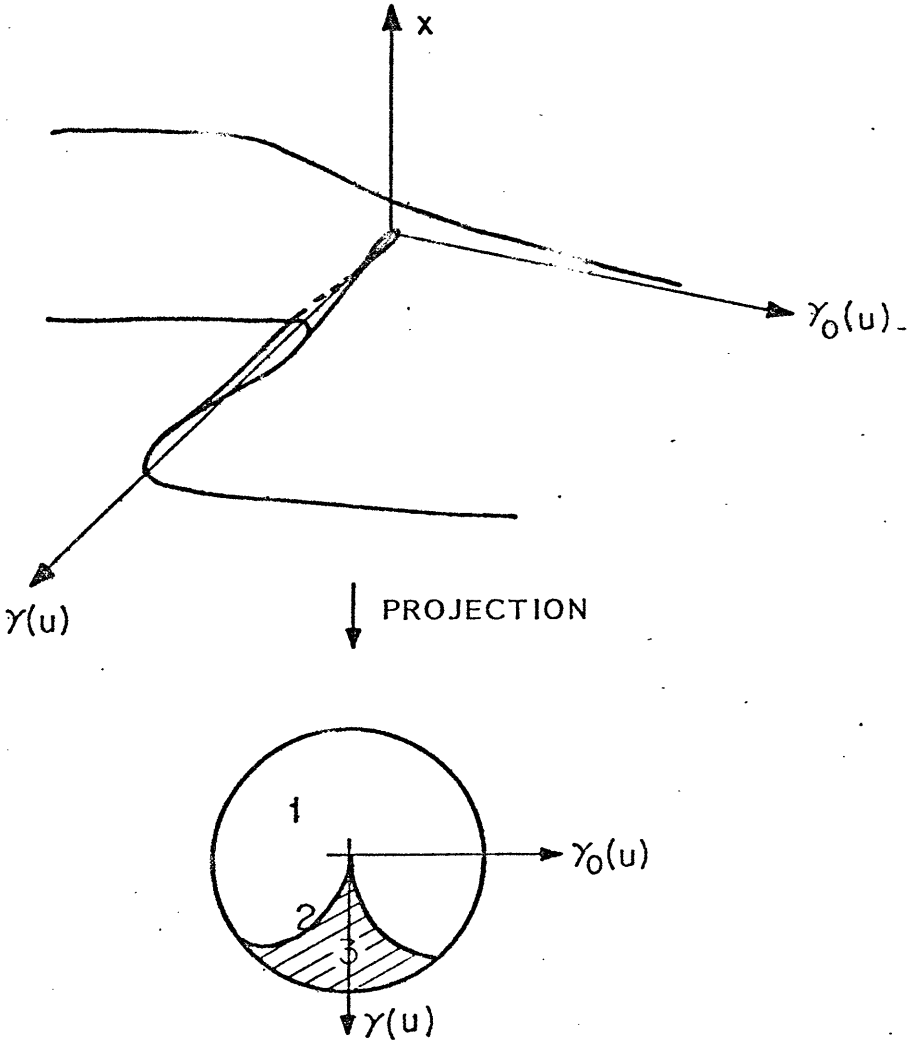

Figure 7. The Cusp Bifurcation

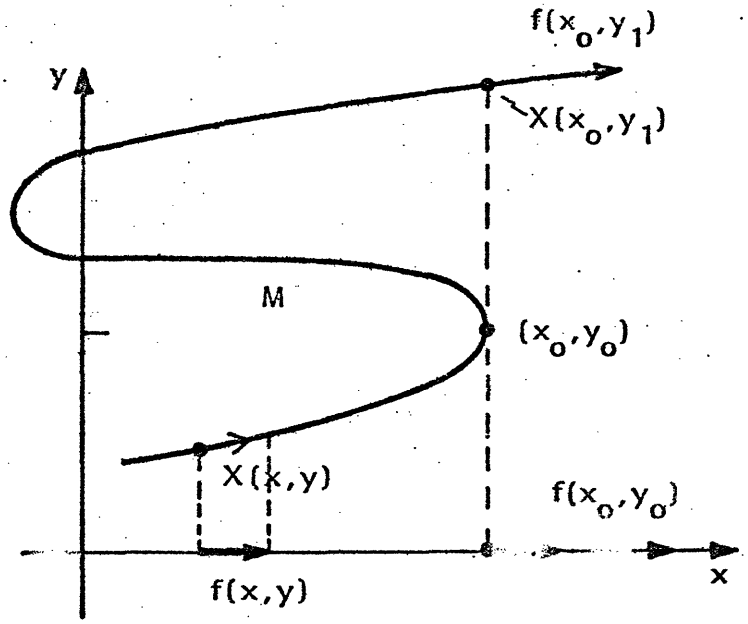

Figure 8. Illustrating Jump Behavior in Constrained Systems

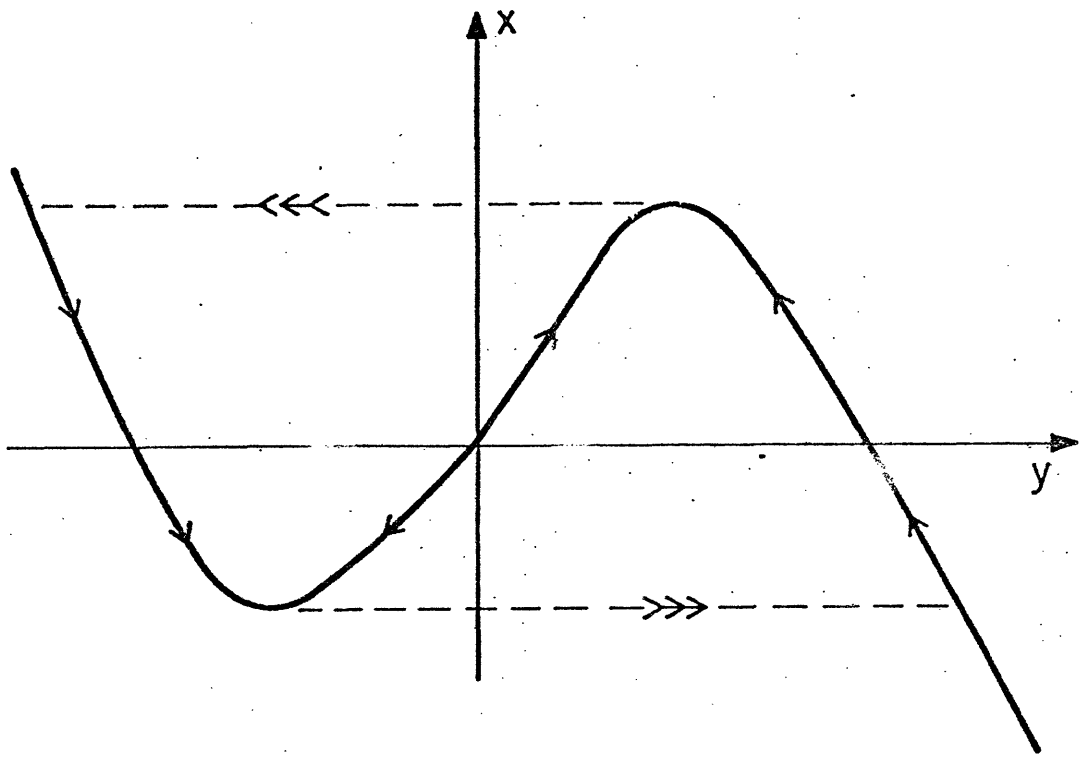

Figure 9. Degenerate Form of the Van Der Pol Oscillator Showing Jump Behavior

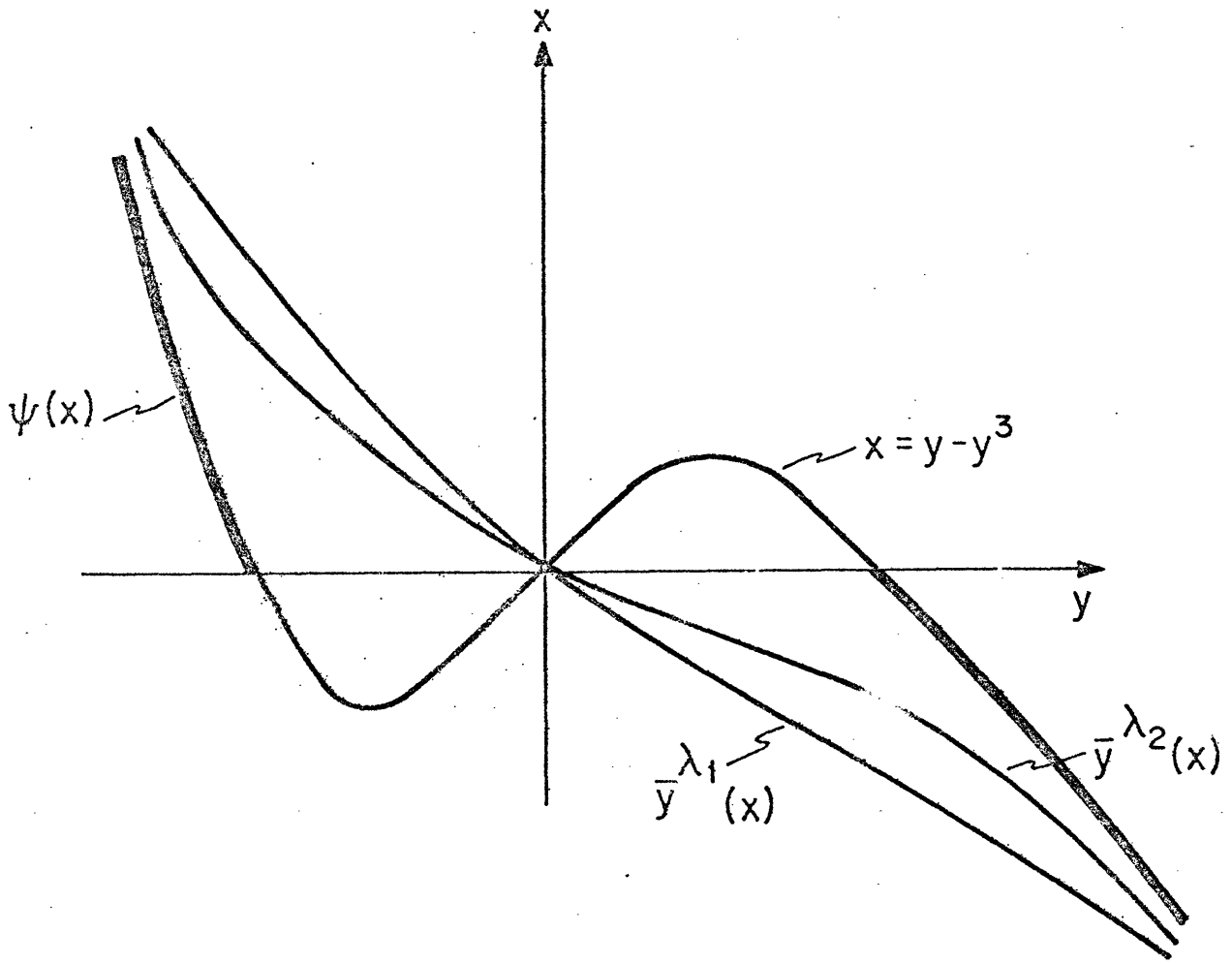

Figure 10. The Drift $\bar{y}^{\lambda}(x)$ for the Limit Diffusion of the Degenerate Van Der Pol Oscillator